Anhang A
Mathematische Grundlagen

In diesem Kapitel werden einige der für das Verständnis dieses Buches benötigten mathematischen Grundlagen kurz eingeführt. Der dargestellte Stoff gehört zu den Grundlagen, die man im Rahmen des Mathematikstudiums im ersten Jahr lernt. Näheres dazu findet man z.B. in Heuser (1988).

A.1 Mengen und Mengenoperationen

Wir verwenden im Folgenden den naiven Mengenbegriff von Cantor. Gemäß diesem ist eine Menge eine Zusammenfassung von wohlunterschiedenen Objekten unseres Denkens und unserer Anschauung. Ist A eine Menge und x ein beliebiges Objekt, so schreiben wir $x \in A$, falls x in der Menge A enthalten ist, und $x \notin A$, falls x nicht in A enthalten ist.

Mengen können explizit durch Aufzählung aller Elemente der Menge definiert werden. Beispiele dafür sind die Menge der natürlichen Zahlen ohne Null

$$\mathbb{N} = \{1,2,3,4,\dots\},$$

die Menge der natürlichen Zahlen mit Null

$$\mathbb{N}_0 = \{0,1,2,3,4,\dots\},$$

oder auch die Menge der ganzen Zahlen

$$\mathbb{Z} = \{\dots,-3,-2,-1,0,1,2,3,\dots\}.$$

Alternativ kann man Mengen durch Angabe von Eigenschaften ihrer Elemente angeben, z.B. die Menge der rationalen Zahlen

$$\mathbb{Q} = \left\{ \frac{p}{q} \quad | \quad p \in \mathbb{Z}, q \in \mathbb{N} \right\}.$$

J. Eckle-Kohler, M. Kohler, *Eine Einführung in die Statistik und ihre Anwendungen,*
DOI 10.1007/978-3-642-15226-9, © Springer-Verlag Berlin Heidelberg 2011

Die Menge $\mathbb{R}$ aller reellen Zahlen besteht aus allen Zahlen, die eine abbrechende oder nicht abbrechende Dezimaldarstellung besitzen, wie z.B. -23.32, $\sqrt{2} = 1.4142\ldots$ oder $\pi = 3.1415\ldots$

Ist A eine Menge reeller Zahlen, so heißt jede Zahl $M \in \mathbb{R}$ mit

$$x \leq M \quad \text{für alle} \quad x \in A$$

obere Schranke von A. Man kann zeigen, dass es immer eine kleinste obere Schranke gibt. Diese wird als *Supremum* von A bezeichnet und wir verwenden dafür die Notation

$$\sup A.$$

Analog heißt jede Zahl $M \in \mathbb{R}$ mit

$$x \geq M \quad \text{für alle} \quad x \in A$$

untere Schranke von A, es gibt eine größte untere Schranke $\inf A$, die als Infimum von A bezeichnet wird.

Ist z.B.

$$A = \{x \in \mathbb{R} : 1 \leq x < 5\}$$

so sind 6, 9 und 15 obere Schranken und -5, 0 und 0.5 untere Schranken von A. Die kleinste obere Schranke ist

$$\sup A = 5,$$

und die größte untere Schranke ist

$$\inf A = 1.$$

Sind A, B zwei Mengen, so können wir daraus neue Mengen bilden durch Vereinigung (d.h. durch Bildung der Menge aller Elemente, die in der ersten oder in der zweiten Menge enthalten sind)

$$A \cup B = \{x \quad | \quad x \in A \text{ oder } x \in B\},$$

durch Durchschnitt (d.h. durch Bildung der Menge aller Elemente, die in der ersten und in der zweiten Menge enthalten sind)

$$A \cap B = \{x \quad | \quad x \in A \text{ und } x \in B\},$$

oder durch Differenzbildung (d.h. durch Bildung der Menge aller Elemente, die in der ersten aber nicht in der zweiten Menge enthalten sind)

$$A \setminus B = \{x \quad | \quad x \in A \text{ und } x \notin B\},$$

vgl. Abbildung A.1.

Die Menge, die kein einziges Element enthält, ist die sog. leere Menge und wird mit dem Symbol $\emptyset$ bezeichnet. Für eine beliebige Menge A gilt immer

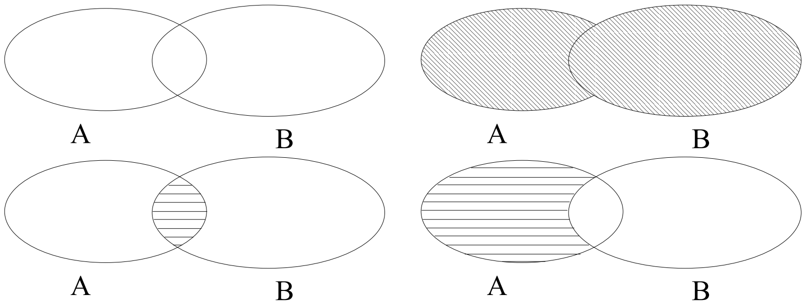

Abb. A.1 Illustration der Vereinigung (oben rechts), des Durchschnitts (unten links) und der Differenz (unten rechts) zweier Mengen.

$$A \cup \emptyset = A, A \cap \emptyset = \emptyset, A \setminus \emptyset = A.$$

Für die Mengenoperationen gelten die folgenden elementaren Beziehungen, wie man sich mit elementaren logischen Umformungen klarmacht:

$$(A_1 \cup A_2) \cap B = (A_1 \cap B) \cup (A_2 \cap B)$$

und

$$(A_1 \cup A_2) \setminus B = (A_1 \setminus B) \cup (A_2 \setminus B).$$

Z.B. liegt x genau dann in der Menge $(A_1 \cup A_2) \cap B$, wenn gilt

$$(x \in A_1 \text{ oder } x \in A_2) \text{ und } x \notin B,$$

was logisch das Gleiche bedeutet wie

$$(x \in A_1 \text{ und } x \notin B) \text{ oder } (x \in A_2 \text{ und } x \notin B).$$

Letzteres wiederum gilt genau dann, wenn x in der Menge $(A_1 \setminus B) \cup (A_2 \setminus B)$ liegt.

Desweiteren gelten die sog. de Morganschen Regeln, wie man sich leicht entweder anschaulich an Abbildung A.2 oder mit elementaren logischen Umformungen klarmacht:

$$A \setminus (B_1 \cup B_2) = (A \setminus B_1) \cap (A \setminus B_2)$$

sowie

$$A \setminus (B_1 \cap B_2) = (A \setminus B_1) \cup (A \setminus B_2).$$

Die obigen Definitionen und Beziehungen lassen sich sofort auf beliebige Anzahlen von Mengen verallgemeinern: Ist z.B. für jedes j aus einer Indexmenge J eine Menge B_j gegeben, so definieren wir

$$\bigcup_{j \in J} B_j = \{x \quad | \quad \text{Für ein } j \in J \text{ gilt } x \in B_j\}$$

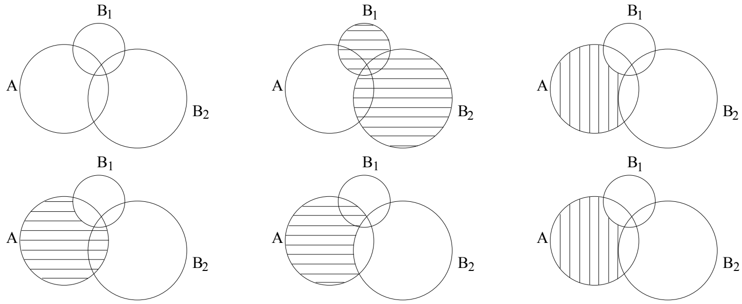

Abb. A.2 Illustration einer der beiden de Morganschen Regeln: Darstellung von drei Mengen A, B_1, B_2 (oben links), sowie $B_1 \cup B_2$ (oben Mitte), $A \setminus (B_1 \cup B_2)$ (oben rechts), $A \setminus B_1$ (unten links), $A \setminus B_2$ (unten Mitte) und $(A \setminus B_1) \cap (A \setminus B_2)$ (unten rechts).

und

$$\bigcap_{j \in J} B_j = \left\{x \quad | \quad \text{Für alle } j \in J \text{ gilt } x \in B_j\right\}$$

und erhalten analog zu oben wieder die de Morganschen Regeln

$$A \setminus \left(\bigcup_{j \in J} B_j\right) = \bigcap_{j \in J} (A \setminus B_j)$$

und

$$A \setminus \left(\bigcap_{j \in J} B_j\right) = \bigcup_{j \in J} (A \setminus B_j).$$

A.2 Das Summenzeichen

In der Statistik bilden wir häufig Summen von reellen Messgrößen. Ist dabei die Anzahl der Summanden groß, oder sind die Werte der Summanden wie in

$$x_1 + x_2 + \cdots + x_n$$

nicht explizit angegeben, so bietet es sich an, für die Summe ein Abkürzung zu verwenden. In der Mathematik verwendet man dafür das griechische Symbol $\sum$ (gesprochen: „sigma“) und schreibt

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

(gesprochen: „Summe der x_i von $i = 1$ bis n“). Entsprechend definieren wir

$$\sum_{i=k}^{n} x_i = x_k + x_{k+1} + \cdots + x_n$$

für $k \leq n$ sowie

$$\sum_{i \in \{2,3,4\}} x_i = x_2 + x_3 + x_4.$$

Bei der letzten Notation nutzen wir aus, dass die Reihenfolge des Aufsummierens (die auf der linken Seite ja nicht eindeutig festgelegt ist) bei endlichen Summen keine Rolle spielt.

Für endliche Summen gelten die üblichen Rechenregeln, z.B. ist

$$\sum_{k=1}^{n} (\alpha \cdot x_k + \beta \cdot y_k) = \alpha \cdot (\sum_{k=1}^{n} x_k) + \beta \cdot (\sum_{k=1}^{n} y_k),$$

wie man sich leicht durch Ausschreiben aller Abkürzungen und Ausmultiplizieren der rechten Seite klarmacht.

Analog sieht man auch die Gültigkeit von sog. Indextransformationen ein, wie sie z.B. in

$$\sum_{i=2}^{n+1} x_i = \sum_{k=1}^{n} x_{k+1}$$

auftritt.

Eine wichtige Summe ist die sog. geometrische Summe

$$\sum_{k=0}^{n} x^k$$

mit $n \in \mathbb{N}$ und $x \in \mathbb{R}$. Diese Summe lässt sich leicht explizit berechnen: Ist nämlich $x = 1$, so summiert man $(n+1)$-mal die Zahl 1 auf und erhält $n+1$. Ist dagegen $x \neq 1$, so gilt

$$(1-x) \cdot \sum_{k=0}^{n} x^k = \sum_{k=0}^{n} x^k - x \cdot \sum_{k=0}^{n} x^k = \sum_{k=0}^{n} x^k - \sum_{k=0}^{n} x^{k+1} = \sum_{k=0}^{n} x^k - \sum_{k=1}^{n+1} x^k = 1 - x^{n+1},$$

woraus

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$$

folgt.

A.3 Folgen und Reihen

In diesem Abschnitt führen wir die Begriffe Folge und Reihe ein und beschreiben einige ihrer wichtigsten Eigenschaften. Wir beginnnen dazu mit

Definition A.1. Eine (reelle) *Folge* ist eine Zuordnung f, die jeder natürlichen Zahl $n \in \mathbb{N}$ in eindeutiger Weise ein Element $f(n) \in \mathbb{R}$ zuordnet. Die Elemente $a_n = f(n)$ heißen *Glieder der Folge* f.

Für Folgen verwenden wir die folgenden beiden Schreibweisen:

$$f = (a_n)_{n\in\mathbb{N}} \quad \text{und} \quad f = (a_n)_n.$$

Beispiel A.1. Beispiele für Folgen sind

a) $f = (a_n)_n$ mit $a_n = 2 + \left(\frac{1}{2}\right)^n$.
b) $g = (b_n)_n$ mit $b_n = \frac{1}{n!} = \frac{1}{n\cdot(n-1)\cdot(n-2)\cdots 1}$.
c) $h = (c_n)_n$ mit $c_n = \sqrt{n}$.

Die Folgenglieder in den Teilen a) und b) des obigen Beispiels nähern sich für wachsendes n immer mehr den Zahlen 2 bzw. 0 an. Falls ein solches Verhalten vorliegt (d.h., falls die Folgenglieder mit wachsendem n immer näher bei einer festen Zahl zu liegen kommen), so spricht man davon, dass die Folge gegen die entsprechende Zahl konvergiert. Dies formalisiert

Definition A.2. Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt *konvergent* gegen $a \in \mathbb{R}$, falls gilt: Für jedes $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt:

$$|a_n - a| < \varepsilon.$$

In diesem Fall heißt a *Grenzwert* der Folge $(a)_{n\in\mathbb{N}}$. Als Schreibweisen verwendet man dafür:

$$\lim_{n\to\infty} a_n = a \quad \text{bzw.} \quad a_n \to a \quad (n \to \infty).$$

Anschaulich bedeutet die obige Definition, dass der Grenzwert a einer Folge die folgende Eigenschaft hat: Für jeden festen Abstand $\varepsilon > 0$ gibt es nur endlich viele Folgenglieder, deren Abstand von a größer oder gleich ε ist.

Bemerkung: Der Grenzwert einer Folge ist – sofern er überhaupt existiert – eindeutig. Sind nämlich a_1 und a_2 zwei verschiedene reelle Zahlen, so gibt es keine Zahl $a \in \mathbb{R}$, die gleichzeitig

$$|a - a_1| < \frac{|a_1 - a_2|}{2} \quad \text{und} \quad |a - a_2| < \frac{|a_1 - a_2|}{2}$$

erfüllt (denn aus der Existenz einer solchen Zahl würde

$$|a_1 - a_2| \leq |a_1 - a| + |a - a_2| < \frac{|a_1 - a_2|}{2} + \frac{|a_1 - a_2|}{2} = |a_1 - a_2|$$

folgen, was nicht gelten kann). Also können nicht gleichzeitig alle bis auf endlich viele Folgenglieder einen Abstand kleiner als $|a_1 - a_2|/2$ zu a_1 und alle bis auf endlich viele Folgenglieder einen Abstand kleiner als $|a_1 - a_2|/2$ zu a_2 haben.

Beispiel A.2. Für die Folge $a_n = (1/2)^n$ gilt

$$a_n \to 0 \quad (n \to \infty).$$

Allgemeiner gilt sogar für jedes beliebige feste $x \in (-1,1)$:

$$x^n \to 0 \quad (n \to \infty).$$

Dagegen konvergiert die Folge $(\sqrt{n})_n$ nicht (da die Folgenglieder beliebig groß werden und damit irgendwann insbesondere einen Abstand größer als Eins von jeder festen Zahl haben).

Es gelten die folgenden Rechenregeln für konvergente Folgen:

Lemma A.1. *Aus*

$$a_n \to a \quad (n \to \infty) \quad \textit{und} \quad b_n \to b \quad (n \to \infty)$$

folgt

$$a_n + b_n \to a + b \quad (n \to \infty),$$
$$a_n - b_n \to a - b \quad (n \to \infty),$$
$$a_n \cdot b_n \to a \cdot b \quad (n \to \infty),$$

sowie im Falle $b \neq 0$ auch

$$\frac{a_n}{b_n} \to \frac{a}{b} \quad (n \to \infty).$$

Sukzessive Anwendung der obigen Rechenregeln ermöglicht die Berechnung vieler komplizierter Grenzwerte durch Zurückführung auf einfache Grenzwerte:

Beispiel A.3. Es gilt

$$1 + \frac{1}{n} \to 1 + 0 = 1 \quad (n \to \infty)$$

und

$$\frac{n^2+2n+1}{2n^2+3} = \frac{1+\frac{2}{n}+\frac{1}{n^2}}{2+\frac{3}{n^2}} \to \frac{1+0+0}{2+0} = \frac{1}{2} \quad (n \to \infty).$$

In Definition A.2 ist der Grenzwert eine reelle Zahl. Es gibt aber auch Folgen, die gegen Unendlich streben. Hierbei schreiben wir

$$a_n \to \infty \quad (n \to \infty),$$

falls für jedes $M \in \mathbb{R}$ ein n_0 existiert mit $a_n > M$ für alle $n \geq n_0$, d.h., falls die Folgenglieder größer werden als jede feste Zahl, sofern der Index nur groß genug ist. Analog schreiben wir

$$a_n \to -\infty \quad (n \to \infty),$$

falls für jedes $M \in \mathbb{R}$ ein n_0 existiert mit $a_n < M$ für alle $n \geq n_0$.

Als Nächstes führen wir den Begriff der (unendlichen) Reihe ein. Dabei handelt es sich um eine Folge, deren Folgenglieder durch Aufsummieren der Glieder einer anderen Folge gebildet werden.

Definition A.3. Sei $(a_n)_n$ eine reelle Folge.
a) Unter der (unendlichen) *Reihe*

$$\sum_{n=1}^{\infty} a_n$$

verstehen wir die Folge (s_n) mit

$$s_n = \sum_{k=1}^{n} a_k \quad \text{für } n \in \mathbb{N}.$$

Dabei heißen die a_n *Glieder* und die s_n *Partialsummen* (Teilsummen) der Reihe.
b) Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *konvergent*, wenn die Folge

$$\left(\sum_{k=1}^{n} a_k\right)_n$$

konvergent ist. In diesem Fall verwenden wir die Schreibweise

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

Man beachte, dass wir in der obigen Definition das gleiche Symbol für zwei verschiedene Dinge verwendet haben:

$$\sum_{n=1}^{\infty} a_n$$

ist sowohl eine Abkürzung für die zugehörige Folge der Partialsummen, als auch für den Grenzwert dieser Folge.

Bemerkung: Unter

$$\sum_{n=0}^{\infty} a_n$$

versteht man die Reihe

$$\sum_{n=1}^{\infty} a_{n-1}.$$

Beispiel A.4. Sei $x \in \mathbb{R}$ fest. Dann heißt

$$\sum_{n=0}^{\infty} x^n$$

geometrische Reihe. Wegen

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} \quad \text{für } x \neq 1$$

und

$$\frac{1-x^{n+1}}{1-x} \to \frac{1-0}{1-x} \quad (n \to \infty) \quad \text{für } |x| < 1$$

gilt für $|x| < 1$:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Für Reihen nichtnegativer Zahlen gilt:

Lemma A.2. *Ist $(a_n)_n$ eine nichtnegative Folge, für die die zugehörige Folge der Partialsummen beschränkt ist, d.h., für die für ein $C \in \mathbb{R}$ gilt:*

$$\left| \sum_{k=1}^{n} a_k \right| \leq C \quad \textit{für alle } n \in \mathbb{N},$$

so ist die Reihe $\sum_{k=1}^{\infty} a_k$ konvergent und der Grenzwert ist unabhängig davon, in welcher Reihenfolge man die Zahlen aufaddiert.

Dieses Lemma hat unmittelbare Anwendung in der Wahrscheinlichkeitstheorie: Ist nämlich $(\Omega, \mathscr{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und sind die Ereignisse $A_1, A_2, A_3 \dots \in \mathscr{A}$ paarweise disjunkt, so sind die Wahrscheinlichkeiten $\mathbf{P}(A_1)$, $\mathbf{P}(A_2)$, $\mathbf{P}(A_3)$ ... nichtnegativ und es gilt

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \dots + \mathbf{P}(A_n) = \mathbf{P}\left(\cup_{k=1}^{n} A_k\right) \leq \mathbf{P}(\Omega) = 1.$$

Anwendung des obigen Lemmas ergibt, dass die Reihe

$$\sum_{k=1}^{\infty} \mathbf{P}(A_k)$$

existiert und der Grenzwert nicht von der Summationsreihenfolge abhängt. Daher sind auch Ausdrücke wie

$$\sum_{k \in \mathbb{N}} \mathbf{P}(A_k),$$

bei denen die Summationsreihenfolge nicht a priori festgelegt ist, wohldefiniert.

Das nächste Lemma formuliert einige Rechenregeln für konvergente Reihen.

Lemma A.3. *Sind $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ zwei konvergente Reihen, so konvergieren auch die Reihen*

$$\sum_{k=1}^{\infty} (a_k + b_k) \quad \textit{und} \quad \sum_{k=1}^{\infty} (a_k - b_k)$$

und es gilt

$$\sum_{k=1}^{\infty}(a_k+b_k)=\sum_{k=1}^{\infty}a_k+\sum_{k=1}^{\infty}b_k \quad \textit{sowie} \quad \sum_{k=1}^{\infty}(a_k-b_k)=\sum_{k=1}^{\infty}a_k-\sum_{k=1}^{\infty}b_k.$$

Obiges Lemma folgt fast unmittelbar aus Lemma A.1. Denn unter den Voraussetzungen des Lemmas gilt z.B.:

$$\begin{aligned}
\sum_{k=1}^{\infty}(a_k+b_k) &\overset{Def.}{=} \lim_{n\to\infty}\sum_{k=1}^{n}(a_k+b_k)\\
&= \lim_{n\to\infty}\left(\sum_{k=1}^{n}a_k+\sum_{k=1}^{n}b_k\right)\\
&\overset{Lemma\ A.1}{=} \lim_{n\to\infty}\sum_{k=1}^{n}a_k+\lim_{n\to\infty}\sum_{k=1}^{n}b_k\\
&\overset{Def.}{=} \sum_{k=1}^{\infty}a_k+\sum_{k=1}^{\infty}b_k.
\end{aligned}$$

A.4 Differentialrechnung

Sei $f:\mathbb{R}\to\mathbb{R}$ eine Funktion (d.h., f weist jeder reellen Zahl $x\in\mathbb{R}$ in eindeutiger Art und Weise eine reelle Zahl $f(x)\in\mathbb{R}$ zu). Im Folgenden wollen wir die Steigung der Funktion in einem Punkt $x_0\in\mathbb{R}$ bestimmen. Dazu betrachten wir eine Gerade, die sowohl durch den Punkt $(x_0,f(x_0))$ geht als auch durch einen weiteren Punkt $(x_1,f(x_1))$. Diese Gerade hat dann die Steigung

$$\frac{f(x_1)-f(x_0)}{x_1-x_0}.$$

Um die Steigung der Funktion f in x_0 zu bestimmen, lassen wir bei der oberen Gerade x_1 gegen x_0 konvergieren (d.h., wir nähern uns mit x_1 immer mehr x_0 an). Sofern sich dann bei den Steigungen der zugehörigen Geraden ein eindeutiger Grenzwert ergibt, bezeichnen wir die Funktion im Punkt x_0 als differenzierbar und betrachten diesen Grenzwert als Steigung von f im Punkt x_0.

Definition A.4. Sei $f:\mathbb{R}\to\mathbb{R}$ eine Funktion.
a) f heißt *differenzierbar* im Punkt $x_0\in\mathbb{R}$, falls eine reelle Zahl b existiert, so dass für jede Folge $(x_n)_n$ mit $x_n\to x_0$ $(n\to\infty)$ und $x_n\neq x_0$ für alle n gilt:

$$\frac{f(x_n)-f(x_0)}{x_n-x_0}\to b \quad (n\to\infty).$$

In diesem Fall heißt b *Ableitung* von f an der Stelle x_0 und wir verwenden die Schreibweise

$$b=f'(x_0).$$

b) f heißt differenzierbar, wenn f für jedes $x_0 \in \mathbb{R}$ differenzierbar im Punkt x_0 ist. In diesem Falle heißt

$$f' : \mathbb{R} \to \mathbb{R} \quad \text{mit} \quad x \mapsto f'(x)$$

Ableitung von f.

Die Ableitungen einiger wichtiger Funktionen sind in Tabelle A.1 angegeben.

Tabelle A.1 Einige Funktionen und ihre Ableitungen.

Funktion	Ableitung
$f(x) = c$ für $c \in \mathbb{R}$ fest	$f'(x) = 0$
$f(x) = x^a$ mit $a \in \mathbb{R} \setminus \{0\}$ fest	$f'(x) = a \cdot x^{a-1}$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = \sin x$	$f'(x) = \cos x$
$f(x) = \cos x$	$f'(x) = -\sin x$
$f(x) = \ln x$	$f'(x) = \frac{1}{x}$

Weitere Ableitungen kann man daraus mit den folgenden drei wichtigen Ableitungsregeln berechnen.

Lemma A.4. *Sind f und g differenzierbar und sind $\alpha, \beta \in \mathbb{R}$, so gilt:*

a) *(Linearität der Ableitung)*

$$(\alpha \cdot f + \beta \cdot g)'(x) = \alpha \cdot f'(x) + \beta \cdot g'(x).$$

b) *(Produktregel)*

$$(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

c) *(Kettenregel)*

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

Beispiel A.5. Wendet man z.B. Teil a) des obigen Lemmas auf

$$h_1(x) = 2 \cdot x^3 + 4 \cdot \sin(x)$$

an, so erhält man

$$h_1'(x) = 2 \cdot 3 \cdot x^2 + 4 \cdot \cos(x).$$

Anwendung von Teil b) auf $h_2(x) = x \cdot \sin(x)$ (mit $f(x) = x$ und $g(x) = \sin x$) liefert

$$h_2'(x) = 1 \cdot \sin x + x \cdot \cos x.$$

Und schließlich erhalten wir mit Teil c) für die Ableitung von $h_3(x) = \sin(2x^2)$ (wobei wir $f(x) = \sin x$ und $g(x) = 2x^2$ setzen)

$$h_3'(x) = \cos(2x^2) \cdot 2 \cdot 2x.$$

Ableitungen lassen sich zur Bestimmung von Extrempunkten einer differenzierbaren Funktion verwenden. Dabei versteht man unter einem Extrempunkt einer Funktion eine Stelle, an der diese Funktion einen maximalen oder minimalen Wert annimmt.

Definition A.5. a) Ein Punkt $z \in \mathbb{R}$ heißt (lokale) *Maximalstelle* einer Funktion $f : \mathbb{R} \to \mathbb{R}$, wenn es ein $\varepsilon > 0$ gibt, so dass

$$f(z) \geq f(x) \quad \text{für alle} \quad x \in (z-\varepsilon, z+\varepsilon).$$

b) Ein Punkt $z \in \mathbb{R}$ heißt (lokale) *Minimalstelle* einer Funktion $f : \mathbb{R} \to \mathbb{R}$, wenn es ein $\varepsilon > 0$ gibt, so dass

$$f(z) \leq f(x) \quad \text{für alle} \quad x \in (z-\varepsilon, z+\varepsilon).$$

c) Ein Punkt $z \in \mathbb{R}$ heißt (lokale) *Extremstelle* einer Funktion $f : \mathbb{R} \to \mathbb{R}$, wenn z (lokale) Maximalstelle oder (lokale) Minimalstelle von f ist.

Man sieht anschaulich leicht ein, dass eine differenzierbare Funktion an einer Extremstelle die Ableitung Null haben muss, denn an einer solchen Stelle muss für die Steigung einer Sekante, im Falle dass sich der zweite Punkt von links oder von rechts gegen die Extremstelle annähert, gelten, dass die Steigung einmal größer oder gleich Null und einmal kleiner oder gleich Null ist. Dies führt auf folgenden Satz:

Satz A.1. *Sei $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion. Ist dann $z \in \mathbb{R}$ eine (lokale) Extremstelle von f, so gilt*

$$f'(z) = 0. \tag{A.1}$$

Die Bedingung (A.1) ist allerdings nur notwendig, aber nicht hinreichend, für das Vorliegen einer Extremstelle. Z.B. erfüllt $f(x) = x^3$ an der Stelle $z = 0$ die Bedingung $f'(z) = 3 \cdot z^2 = 0$, wegen

$$f(x) < 0 = f(0) \quad \text{für } x < 0 \quad \text{und} \quad f(x) > 0 = f(0) \quad \text{für } x > 0$$

hat f an der Stelle Null aber keine (lokale) Extremstelle.

Vereinzelt betrachten wir auch reelle Funktionen mehrerer Variablen. Ist z.B. $f : \mathbb{R}^2 \to \mathbb{R}$ eine reelle Funktion zweier Variablen, d.h., f weist jedem Paar $(x,y) \in \mathbb{R}^2$ eine reelle Zahl $f(x,y)$ zu, so ist für jedes feste $y \in \mathbb{R}$

$$x \mapsto f(x,y)$$

eine reelle Funktion einer Variablen. Ist diese im obigen Sinne differenzierbar, so verwenden wir die Schreibweise

$$\frac{\partial f}{\partial x}(x,y)$$

für ihre Ableitung. Analog ist für jedes feste $x \in \mathbb{R}$

$$y \mapsto f(x,y)$$

eine reelle Funktion einer Variablen, deren Ableitung wir mit

$$\frac{\partial f}{\partial y}(x,y)$$

bezeichnen.

A.5 Integralrechnung

Seien $a,b \in \mathbb{R}$ mit $a < b$ und sei $f : [a,b] \to \mathbb{R}$ eine auf $[a,b]$ definierte reellwertige Funktion. Bei der Integralrechnung interessiert man sich für den Flächeninhalt zwischen f und der x-Achse im Intervall $[a,b]$.

Dieser lässt sich einfach berechnen, falls f stückweise konstant ist: Ist nämlich f gegeben durch

$$f(x) = \sum_{i=1}^{n} \alpha_i \cdot 1_{[x_{i-1},x_i)}(x) \tag{A.2}$$

für $n \in \mathbb{N}$, $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ und

$$a = x_0 < x_1 < \cdots < x_n = b,$$

so können wir diesen berechnen zu

$$Int_{[a,b]}(f) = \sum_{i=1}^{n} \alpha_i \cdot (x_i - x_{i-1}). \tag{A.3}$$

Sofern nicht alle α_i nichtnegativ sind, handelt es sich dabei um einen orientierten Flächeninhalt.

Durch (A.3) wird ein Integral für die Menge aller Funktionen der Bauart (A.2) definiert. Die Menge aller dieser sog. Treppenfunktionen bezeichnen wir im Folgenden mit $T([a,b])$.

Der obige Integralbegriff kann auf folgende Art auf allgemeine Funktionen erweitert werden:

Definition A.6. Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt (Riemann-) *integrierbar* über $[a,b]$, falls zu jedem $\varepsilon > 0$ Treppenfunktionen $f_1, f_2 \in T([a,b])$ existieren mit

$$f_1(x) \leq f(x) \leq f_2(x) \quad \text{für alle } x \in [a,b]$$

und

$$Int_{[a,b]}(f_2) - Int_{[a,b]}(f_1) < \varepsilon.$$

In diesem Falle heißt

$$\int_a^b f(x)\,dx = \sup\left\{Int_{[a,b]}(g) \quad : \quad g \in T([a,b]) \text{ mit } g(x) \leq f(x) \text{ für alle } x \in [a,b]\right\}$$

(Riemann-) *Integral* von f.

Die Berechnung von Integralen erfolgt mit Hilfe von

Satz A.2. *(Hauptsatz der Differential- und Integralrechnung)*
Ist $F : [a,b] \to \mathbb{R}$ Stammfunktion von $f : [a,b] \to \mathbb{R}$, d.h., ist F differenzierbar mit

$$F'(x) = f(x) \quad \textit{für alle } x \in [a,b],$$

und ist f stetig auf $[a,b]$, so gilt

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Beispiel A.6. $F(x) = -e^{-x}$ ist die Stammfunktion zu $f(x) = e^{-x}$ und daher gilt

$$\int_a^b e^{-x}dx = -e^{-x}\Big|_{x=a}^{b} = -e^{-b} - (-e^{-a}) = -e^{-b} + e^{-a}.$$

Wichtige Rechenregeln bei der Berechnung von Integralen beschreibt der nächste Satz.

Satz A.3. a) *(Linearität) Für Funktionen f und g und für $\alpha, \beta \in \mathbb{R}$ gilt*

$$\int_a^b (\alpha \cdot f + \beta \cdot g)(x)\,dx = \alpha \cdot \int_a^b f(x)\,dx + \beta \cdot \int_a^b g(x)\,dx.$$

b) *Für eine Funktion f und $a < b < c$ gilt:*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

c) *(Partielle Integration) Sind f und g stetig differenzierbare Funktionen, und sind $a, b \in \mathbb{R}$ mit $a < b$, so gilt*

$$\int_a^b f(x) \cdot g'(x)\,dx = f(x) \cdot g(x)\Big|_{x=a}^{b} - \int_a^b f'(x) \cdot g(x)\,dx.$$

Die Formel von der partiellen Integration folgt unmittelbar aus der Produktregel für Ableitungen, der Linearität des Integrals und dem Hauptsatz der Differential- und Integralrechnung, denn:

$$\begin{aligned} f(x) \cdot g(x)\Big|_{x=a}^{b} &= \int_a^b (f \cdot g)'(x)\,dx = \int_a^b (f'(x) \cdot g(x) + f(x) \cdot g'(x))\,dx \\ &= \int_a^b f'(x) \cdot g(x)\,dx + \int_a^b f(x) \cdot g'(x)\,dx. \end{aligned}$$

Beispiel A.7. Setzt man $f(x) = x$ und $g'(x) = e^{-x}$ (was $g(x) = -e^{-x}$ impliziert), so folgt aus der Produktregel

$$\begin{aligned}\int_a^b x \cdot e^{-x} dx &= -x \cdot e^{-x}\Big|_{x=a}^b - \int_a^b 1 \cdot (-e^{-x}) dx \\ &= -b \cdot e^{-b} + a \cdot e^{-a} - e^{-x}\Big|_{x=a}^b \\ &= -b \cdot e^{-b} + a \cdot e^{-a} - e^{-b} + e^{-a} \\ &= -(b+1) \cdot e^{-b} + (a+1) \cdot e^{-a}.\end{aligned}$$

Definition A.7. Für $a, b \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ setzen wir – sofern existent –

$$\int_a^\infty f(x)\, dx = \lim_{\beta \to \infty} \int_a^\beta f(x)\, dx,$$

$$\int_{-\infty}^b f(x)\, dx = \lim_{\alpha \to -\infty} \int_\alpha^b f(x)\, dx$$

und

$$\int_{-\infty}^\infty f(x)\, dx = \int_{-\infty}^0 f(x)\, dx + \int_0^\infty f(x)\, dx.$$

Die obigen Integrale heißen *uneigentliche Integrale*.

In obiger Definition schreiben wir für $a \in \mathbb{R}$ und $F : \mathbb{R} \to \mathbb{R}$

$$\lim_{\beta \to \infty} F(\beta) = a,$$

sofern für jede Folge $(\beta_n)_n$ mit $\beta_n \to \infty$ $(n \to \infty)$ gilt

$$F(\beta_n) \to a \quad (n \to \infty).$$

Analog ist

$$\lim_{\alpha \to -\infty} F(\alpha)$$

definiert.

Beispiel A.8.

$$\begin{aligned}\int_0^\infty x \cdot e^{-x} dx &= \lim_{\beta \to \infty} \int_0^\beta x \cdot e^{-x} dx \overset{s.o.}{=} \lim_{\beta \to \infty} \left(-(\beta+1) \cdot e^{-\beta} + (0+1) \cdot e^{-0} \right) \\ &= 0 + 1 = 1.\end{aligned}$$

Anmerkungen

Anmerkungen zu Kapitel 1

1. Eine solche Studie, bei der 1074 Männer und 1013 Frauen im Alter zwischen 45 und 64 Jahren in Augsburg über einen Zeitraum von 13 Jahren (1984-1997) beobachtet wurden, ist in Heidrich et al. (2003) beschrieben.
2. Die Daten aus diesem Beispiel stammen aus Ebrahim et al. (2002). Die Auswertung dort versucht auch konfundierende Faktoren (vgl. Kapitel 2) zu berücksichtigen und kommt dabei zu einem ähnlichen Ergebnis wie bei direkter Betrachtung der Zahlen.
3. Eine ausführliche Beschreibung der Challenger-Katastrophe und potentieller Hintergründe findet man in Rogers et al. (1986) sowie bei Wikipedia unter
   http://en.wikipedia.org/wiki/Space_Shuttle_Challenger_disaster
   Die Darstellung in diesem Abschnitt orientiert sich unter anderem an Tufte (1997).
4. Eine ausführliche Beschreibung des Ablaufs der Amerikanischen Präsidentschaftswahl 2000 findet man bei Wikipedia unter
   http://en.wikipedia.org/wiki/United_States_presidential_election_2000
5. Mehr zu den Prozessen in Florida findet man im Internet unter
   http://www.floridasupremecourt.org/pub_info/election/index.shtml
6. Näheres zum Parkfield Earthquake Experiment findet man z.B. unter
   http://earthquake.usgs.gov/research/parkfield/overview.php
7. Fragestellungen dieser Art werden z.B. in der Geodäsie untersucht, siehe z.B. Grafarend (2003).
8. Die Biostatistik ist eines der zur Zeit vielversprechendsten Anwendungsgebiete der Statistik. Eine Einführung in aktuelle Fragestellungen im Zusammenhang mit DNA Microarrays findet sich z.B. in Amaratunga und Cabreva (2003).
9. Eine Einführung in die Schadenversicherungsmathematik, die die zugrundeliegenden mathematischen Probleme in realen Anwendungen schön beschreibt, findet man z.B. in Mack (2002).

J. Eckle-Kohler, M. Kohler, *Eine Einführung in die Statistik und ihre Anwendungen,*
DOI 10.1007/978-3-642-15226-9, © Springer-Verlag Berlin Heidelberg 2011

10. Für eine allgemeine und sehr umfassende Einführung in die Finanzmathematik sei auf Hull (2006) verwiesen. Näheres zu statistischen Verfahren zur Bewertung des Risikos von Investitionen in Kapitalanlagen findet man z.B. in Franke, Härdle and Hafner (2007).

Anmerkungen zu Kapitel 2

1. Die Darstellung der Entwicklung und Überprüfung der Wirksamkeit von Tamiflu orientiert sich an Schneider (2002).
2. Im Rahmen des sog. Mikrozensus (regelmäßige Befragung von 1% der Bevölkerung in Deutschland, zum Teil mit Auskunftspflicht) führt das statistische Bundesamt auch Befragungen zu den Rauchgewohnheiten der Bevölkerung durch. Im Jahr 2005 haben dabei 32.2% der befragten Männer und 22.4% der befragten Frauen angegeben, regelmäßig zu rauchen. Von den 20-25-Jährigen zählten sich 40.3%, von den 60-65-Jährigen aber nur 18.8% zu den Rauchern. (Quelle: Statistisches Bundesamt. Mikrozensus – Fragen zur Gesundheit – Rauchgewohnheiten der Bevölkerung 2005.)
3. Details dieser Studie findet man in Der, Butty und Derry (2006).
4. Mehr zur Nurses Health Study findet man z.B. bei Wikepedia unter http://de.wikipedia.org/wiki/Nurses´_Health_Study
5. Diese Daten stammen aus Heart Protection Study Collaborative Group (2002).
6. Dieses Beispiel stammt aus Singh et al. (2002)
7. Siehe Gadhia et al. (2003).
8. Mehr dazu findet man unter http://www.emf-forschungsprogramm.de/forschung/biologie/biologie_abges/bio_105.html
9. Diese Studie ist in Weinstein et al. (2006) beschrieben.
10. Näheres dazu findet man z.B. bei SPIEGEL ONLINE unter http://www.spiegel.de/wissenschaft/mensch/0,1518,406077,00.html
11. Dieses Beispiel wird unter http://www.zeit.de/2004/32/Pinguine bei ZEIT ONLINE beschrieben.
12. Die Umfrageergebnisse sind zu finden unter http://www.bundestagswahl-2009.de/wahlumfragen/
13. Eine einfache Methode dafür ist die Befragung der Person im Haushalt, die zuletzt Geburtstag hatte (sog. Last-Birthday-Verfahren).
14. Mehr zur Gewichtung der Angaben bei Wahlumfragen findet man z.B. in Roth (2008).

Anmerkungen zu Kapitel 3

1. Die in diesem Beispiel verwendeten Daten wurden dem vom statistischen Bundesamt veröffentlichten Statistischen Jahrbuch 2003 entnommen.
2. Der Kern-Dichteschätzer geht auf Rosenblatt (1956) und Parzen (1962) zurück.
3. Der Kern-Dichteschätzer wird hier im Sinne der explorativen Statistik auf diskrete Variablen angewendet, obwohl a priori klar ist, dass solche Daten niemals durch eine Dichte exakt beschrieben werden können. Allerdings können – wie das Beispiel zeigt – durch Anwenden des Kerndichteschätzers Aussagen über das globale Verhalten von relativen Häufigkeiten gewonnen werden.
4. Die Daten wurden analog zu Angaben in einem so genannten CAMPUS-file des Forschungsdatenzentrums des statistischen Bundesamtes gebildet, das entsprechende CAMPUS-file findet man unter

   http://www.forschungsdatenzentrum.de/

5. Eurostat ist die Statistik-Behörde der Europäischen Union. Ihre Homepage findet man im Internet unter

   http://epp.eurostat.ec.europa.eu/

   Dort sind auch die Daten erhältlich, die in diesem Beispiel verwendet wurden.
6. Eine umfassende mathematische Theorie zur nichtparametrischen Regressionsschätzung findet man in Györfi et al. (2002).

Anmerkungen zu Kapitel 4

1. Gesprochen: „omega“.
2. Hierbei handelt es sich um das kleine „omega“.
3. Die sog. Fakultät von n:
   $$n! = n \cdot (n-1) \cdot \cdots \cdot 1$$
   wird gesprochen als „n Fakultät“.
4. Der sog. Binomialkoeffizient
   $$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$
   wird gesprochen als „n über k“.
5. Gesprochen: „sigma“-Additivität.
6. Dass es kein Wahrscheinlichkeitsmaß
   $$\mathbf{P} : \mathscr{P}([0,5]) \to \mathbb{R}$$
   gibt, das eine Gleichverteilung auf $[0,5]$ beschreibt, kann man sich auch folgendermaßen klarmachen:

Angenommen, $\mathbf{P} : \mathscr{P}([0,5]) \to \mathbb{R}$ ist ein Wahrscheinlichkeitsmaß, das eine Gleichverteilung auf $[0,5]$ beschreibt. Dann sollte sich der Wert von $\mathbf{P}(A)$ nicht ändern, wenn man die Menge A zyklisch verschiebt. Das zyklische Verschieben eines Elementes $\omega \in [0,5]$ um eine Distanz $x \in [0,5]$ beschreiben wir durch

$$x \oplus \omega = \begin{cases} \omega + x & \text{falls } \omega + x \leq 5, \\ \omega + x - 5 & \text{falls } 5 < \omega + x \leq 10 \end{cases}$$

Mit Hilfe dieser Operation können wir durch

$$x \oplus A = \{x \oplus \omega \quad : \quad \omega \in A\}$$

die Verschiebung einer Menge $A \subseteq [0,5]$ definieren. Die obige Eigenschaft der Gleichverteilung lässt sich dann durch

$$\mathbf{P}(x \oplus A) = \mathbf{P}(A) \quad \text{für alle } x \in [0,5], A \subseteq [0,5] \tag{A.4}$$

formalisieren. Wir zeigen im Folgenden, dass ein Wahrscheinlichkeitsmaß

$$\mathbf{P} : \mathscr{P}([0,5]) \to \mathbb{R}$$

diese Eigenschaft nicht haben kann.
Dazu setzen wir $Q := [0,5] \cap \mathbb{Q}$ und definieren eine Äquivalenzrelation $\sim$ durch

$$\omega_1 \sim \omega_2 \quad \Leftrightarrow \quad \exists p \in Q : \omega_2 = \omega_1 \oplus q.$$

Die zugehörigen Äquivalenzklassen seien mit

$$[\omega] = \{\tilde{\omega} \in [0,5] \,:\, \tilde{\omega} \sim \omega\}$$

bezeichnet.
Trivialerweise gilt

$$[0,5] = \bigcup_{\omega \in [0,5]} [\omega].$$

Durch Anwendung des Auswahlaxioms können wir eine Teilmenge $R \subseteq [0,5]$ konstruieren mit der Eigenschaft, dass für jedes $\omega \in [0,5]$ genau ein $r \in R$ existiert mit $\omega \in [r]$. Insbesondere gilt dann

$$q_1 \oplus R \cap q_2 \oplus R = \emptyset$$

für alle $q_1, q_2 \in Q$ mit $q_1 \neq q_2$.
Wegen

$$[0,5] = \bigcup_{q \in Q} q \oplus R$$

folgt daraus aber mit der σ-Additivität von $\mathbf{P}$:

$$1 = \mathbf{P}\left(\bigcup_{q \in Q} q \oplus R\right) = \sum_{q \in Q} \mathbf{P}(q \oplus R).$$

Bei einer Gleichverteilung müssen aber die unendlich vielen Summanden rechts wegen (A.4) alle den gleichen Wert haben, der dann aber nicht zu Eins aufsummieren kann.

7. Genauer gilt:
Lemma.
a) $\mathscr{A}_\alpha$ σ-Algebra in Ω $(\alpha \in I \neq \emptyset) \Rightarrow \bigcap_{\alpha \in I} \mathscr{A}_\alpha$ σ-Algebra in Ω.
b) $\mathscr{E}$ Mengensystem in Ω. Dann existiert eine „kleinste" σ-Algebra, die $\mathscr{E}$ umfasst, d. h. es existiert eine σ-Algebra, die $\mathscr{E}$ umfasst und in allen σ-Algebren enthalten ist, die $\mathscr{E}$ umfassen.
Beweis. a) (i) $\emptyset, \Omega \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha$, da $\emptyset, \Omega \in \mathscr{A}_\alpha$ für alle $\alpha \in I$ und $I \neq \emptyset$.
(ii) $A \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha \Rightarrow A \in \mathscr{A}_\alpha$ für alle $\alpha \in I \Rightarrow A^c \in \mathscr{A}_\alpha$ für alle $\alpha \in I$
$\Rightarrow A^c \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha$
(iii) $A, B \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha \Rightarrow A, B \in \mathscr{A}_\alpha$ für alle $\alpha \in I \Rightarrow A \cup B, A \cap B, A \setminus B \in \mathscr{A}_\alpha$ für alle $\alpha \in I \Rightarrow A \cup B, A \cap B, A \setminus B \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha$
(iv) $A_n \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha$ für alle $n \in \mathbb{N} \Rightarrow A_n \in \mathscr{A}_\alpha$ für alle $\alpha \in I$, $n \in \mathbb{N}$
$\Rightarrow \cup_{n \in \mathbb{N}} A_n \in \mathscr{A}_\alpha$ für alle $\alpha \in I \Rightarrow \cup_{n \in \mathbb{N}} A_n \in \bigcap_{\alpha \in I} \mathscr{A}_\alpha$
b) Setze $I := \{\mathscr{A} \subseteq \mathscr{P}(\Omega) : \mathscr{A}$ σ-Algebra mit $\mathscr{E} \subseteq \mathscr{A}\}$
Dann gilt $I \neq \emptyset$, da $\mathscr{P}(\Omega) \in I$.
Nun ist aber $\bigcap_{\mathscr{A} \in I} \mathscr{A}$ eine σ-Algebra (siehe a)), die $\mathscr{E}$ umfasst (nach Definition), und die in jeder σ-Algebra enthalten ist, die $\mathscr{E}$ umfasst (da jede solche σ-Algebra in obigem Schnitt auftaucht). $\square$
8. Siehe z.B. Satz 6.4 in Bauer (1992). Trotz dieser schönen Eigenschaften der Borelschen σ-Algebra stimmt diese aber nicht mit $\mathscr{P}(\mathbb{R})$ überein, siehe z.B. Satz 8.6 in Bauer (1992).
9. Die Existenz eines solchen Maßes folgt aus dem sog. Fortsetzungssatz von Caratheodory, siehe z.B. Satz 5.1 in Bauer (1992).
10. Die Regel von de l'Hospital besagt unter anderem, dass für differenzierbare Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ mit $f(0) = 0 = g(0)$ gilt

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \frac{f'(x)}{g'(x)},$$

sofern der Limes auf der rechten Seite existiert (siehe z.B. Satz 50.1 in Heuser (1988)).
11. Dies folgt durch Anwendung des sog. Satzes von der monotonen Konvergenz, siehe z.B. Satz 11.4 in Bauer (1992).
12. Ist nämlich $(\varepsilon_n)_{n \in \mathbb{N}}$ eine monoton fallende Nullfolge, so gilt

$$\lim_{n \to \infty} \int_{(x-\varepsilon_n, x+\varepsilon_n)} f(u)\, du = \lim_{n \to \infty} \int f(u) \cdot 1_{(x-\varepsilon_n, x+\varepsilon_n)}(u)\, du = \int 0\, du = 0,$$

wobei wir ausgenutzt haben, dass wir nach dem Satz von der majorisierten Konvergenz (siehe z.B. Satz 15.6 in Bauer (1992)) beim letzten Integral den Grenzwert mit dem Integral vertauschen dürfen und dass für $u \neq x$ gilt

$$\lim_{n\to\infty} f(u)\cdot 1_{(x-\varepsilon_n,x+\varepsilon_n)}(u) = f(u)\cdot 0 = 0.$$

13. Diese Beziehung folgt z.B. mit Hilfe von mehrdimensionaler Integralrechnung wie folgt:

$$\begin{aligned}\int_{\mathbb{R}} e^{-x^2/2}\,dx\cdot\int_{\mathbb{R}} e^{-y^2/2}\,dy &= \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} d(x,y)\\ &= \int_0^{2\pi}\int_0^\infty r\cdot e^{-r^2/2}\,dr\,d\theta\\ &= 2\pi\cdot\int_0^\infty r\cdot e^{-r^2/2}\,dr\\ &= 2\pi\cdot(-e^{-r^2/2})\Big|_{r=0}^{\infty} = 2\pi.\end{aligned}$$

14. Die Zahlenangaben in diesem Beispiel stammen aus Hildebrandt (1998).
15. Angaben zur Genauigkeit des Triple-Tests findet man im Internet z.B. unter

http://www.humangenetik.gwdg.de/Seiten/Infoblatt%20Triple.html

16. Bei den hier angegebenen Werten für die Wahrscheinlichkeit, ein Kind mit Down-Syndrom zu bekommen, erscheint der Triple-Test im Rahmen der allgemeinen Vorsorgeuntersuchungen für werdende Mütter aller Altersstufen (und in dieser Form wird er bei privatversicherten Frauen heutzutage eingesetzt) keineswegs sinnvoll. Allerdings bekommt man bei Frauen über 35 andere Werte, da bei diesen der Wert von $\mathbf{P}(A)$ höher ausfällt und damit auch die durch $\mathbf{P}(A|B)$ gegebene Aussagekraft des positiven Testergebnisses steigt.

Anmerkungen zu Kapitel 5

1. Man kann z.B. zeigen, dass jede stetige Abbildung $\mathscr{B}-\mathscr{B}$-messbar ist, dass punktweise Grenzwerte $\mathscr{B}-\mathscr{B}$-messbarer Abbildungen selbst immer $\mathscr{B}-\mathscr{B}$-messbar sind, dass mit zwei Abbildungen auch ihre Summe bzw. ihre Differenz bzw. ihr Produkt bzw. ihr Quotient (sofern definiert) $\mathscr{B}-\mathscr{B}$-messbar sind, und dass die Verkettung $\mathscr{B}-\mathscr{B}$-messbarer Abbildungen eine $\mathscr{B}-\mathscr{B}$-messbar Abbildung ergibt. Siehe z.B. §9 in Bauer (1992).
2. Siehe z.B. Satz 6.5 in Bauer (1992).
3. Nach Satz 6.5 in Bauer (1992) gehört zu F ein Wahrscheinlichkeitsmaß mit Verteilungsfunktion F. Die gesuchte Zufallsvariable erhält man daraus durch Konstruktion einer Zufallsvariablen, deren Verteilung mit diesem Wahrscheinlichkeitsmaß übereinstimmt.
4. Denn gilt

$$\lim_{x\to\infty} F(x) \neq 1,$$

so existiert ein $\varepsilon > 0$ und eine monoton wachsende Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \to \infty$ $(n \to \infty)$ und

$$|F(x_n) - 1| \geq \varepsilon \quad \text{für alle } n \in \mathbb{N}.$$

Letzteres impliziert aber, dass $F(x_n)$ nicht gegen Eins konvergieren kann.

5. Denn für $B \in \mathscr{B}$ gilt wegen der $\mathscr{B} - \mathscr{B}$-Messbarkeit von h_i:

$$h_i^{-1}(B) \in \mathscr{B}.$$

Wegen

$$(h_i(X_i))^{-1}(B) = (h_i \circ X_i)^{-1}(B) = X_i^{-1}(h_i^{-1}(B))$$

folgt daraus aber

$$(h_i(X_i))^{-1}(B) = X_i^{-1}(h_i^{-1}(B)) \in \mathscr{A},$$

da X_i Zufallsvariable ist.

6. Siehe z.B. Satz 9.6 in Bauer (1991).
7. Ist nämlich $f(\bar{\omega}) \leq n$, so gilt

$$|f_n(\bar{\omega}) - f(\bar{\omega})| \leq \frac{1}{2^n},$$

was

$$f_n(\bar{\omega}) \to f(\bar{\omega}) \quad \text{für alle} \quad \bar{\omega} \in \Omega$$

impliziert. Weiter gilt für alle $\bar{\omega} \in \Omega$ und $n \in \mathbb{N}$

$$f_n(\bar{\omega}) \leq f_{n+1}(\bar{\omega}),$$

denn ist

$$\frac{k}{2^n} \leq f(\bar{\omega}) < \frac{k+1}{2^n}$$

so gilt auch

$$\frac{2k}{2^{n+1}} \leq f(\bar{\omega}) < \frac{2k+2}{2^{n+1}},$$

und wählt man nun l so, dass

$$\frac{l}{2^{n+1}} \leq f(\bar{\omega}) < \frac{l+1}{2^{n+1}},$$

gilt, so folgt

$$f_{n+1}(\bar{\omega}) = \frac{l}{2^{n+1}} \geq \frac{2k}{2^{n+1}} = \frac{k}{2^n} = f_n(\bar{\omega}).$$

Darüberhinaus gilt im Falle $f(\bar{\omega}) \geq n$ wegen

$$n = \frac{n \cdot 2^{n+1}}{2^{n+1}} \quad \text{und} \quad n \cdot 2^{n+1} \leq (n+1) \cdot 2^{n+1} - 1$$

ebenfalls

$$f_{n+1}(\bar{\omega}) \geq \frac{n \cdot 2^{n+1}}{2^{n+1}} = n = f_n(\bar{\omega}).$$

8. Siehe z.B. Korollar 11.2 in Bauer (1992).
9. Z.B. ist f^+ messbar, da f^+ Verkettung von f mit der stetigen Abbildung

$$u \mapsto \max\{u, 0\}$$

ist, stetige Abbildungen $\mathscr{B} - \mathscr{B}$-messbar sind (siehe z.B. §7 in Bauer (1992)) und die Verkettung messbarer Abbildungen eine messbare Abbildung ergibt (vgl. Anmerkung 1 oben).
10. Dies folgt aus dem Satz von der monotonen Konvergenz, siehe z.B. Satz 11.4 in Bauer (1992).
11. Einen Beweis findet man z.B. in §8 von Bauer (1991).
12. Es gelte $Z_n \to Z$ $f.s.$ Sei $\varepsilon > 0$ beliebig. Dann ist

$$\lim_{n\to\infty} \mathbf{P}[|Z_n - Z| > \varepsilon] = \lim_{n\to\infty} \int 1_{\{\bar{\omega}\in\Omega : |Z_n(\bar{\omega}) - Z(\bar{\omega})| > \varepsilon\}}(\omega)\mathbf{P}(d\omega).$$

Der Integrand konvergiert mit Wahrscheinlichkeit Eins gegen Null, denn für jedes $\omega \in \Omega$ mit

$$Z_n(\omega) \to Z(\omega) \quad (n \to \infty)$$

gilt

$$1_{\{\bar{\omega}\in\Omega : |Z_n(\bar{\omega}) - Z(\bar{\omega})| > \varepsilon\}}(\omega) = 0$$

für n genügend groß, was

$$\lim_{n\to\infty} 1_{\{\bar{\omega}\in\Omega : |Z_n(\bar{\omega}) - Z(\bar{\omega})| > \varepsilon\}}(\omega) = 0$$

impliziert. Mit dem Satz von der majorisierten Konvergenz (siehe z.B. Satz 15.6 in Bauer (1992)) folgt daraus

$$\lim_{n\to\infty} \int 1_{\{\bar{\omega}\in\Omega : |Z_n(\bar{\omega}) - Z(\bar{\omega})| > \varepsilon\}}(\omega)\mathbf{P}(d\omega) = \int 0\, \mathbf{P}(d\omega) = 0.$$

Also gilt

$$\lim_{n\to\infty} \mathbf{P}[|Z_n - Z| > \varepsilon] = 0.$$

13. Zum Beweis beachte man, dass aus $X_n(\omega) \to X(\omega)$ $(n \to \infty)$ und $Y_n(\omega) \to Y(\omega)$ $(n \to \infty)$ immer $\alpha \cdot X_n(\omega) + \beta \cdot Y_n(\omega) \to \alpha \cdot X(\omega) + \beta \cdot Y(\omega)$ $(n \to \infty)$ folgt. Durch Kontraposition erhält man daraus, dass

$$\alpha \cdot X_n(\omega) + \beta \cdot Y_n(\omega) \nrightarrow \alpha \cdot X(\omega) + \beta \cdot Y(\omega) \quad (n \to \infty)$$

die Beziehung

$$X_n(\omega) \nrightarrow X(\omega) \quad (n \to \infty) \quad \text{oder} \quad Y_n(\omega) \nrightarrow Y(\omega) \quad (n \to \infty)$$

zur Folge hat. Dies zeigt

$$\begin{aligned}&[\alpha \cdot X_n + \beta \cdot Y_n \nrightarrow \alpha \cdot X + \beta \cdot Y \ (n \to \infty)] \\ &\subseteq [X_n \nrightarrow X \ (n \to \infty)] \cup [Y_n \nrightarrow Y \ (n \to \infty)].\end{aligned}$$

Aus $X_n \to X$ $f.s.$ und $Y_n \to Y$ $f.s.$ folgt daher

$$\begin{aligned}&\mathbf{P}[\alpha \cdot X_n + \beta \cdot Y_n \nrightarrow \alpha \cdot X + \beta \cdot Y \ (n \to \infty)] \\ &\leq \mathbf{P}[X_n \nrightarrow X \ (n \to \infty)] + \mathbf{P}[Y_n \nrightarrow Y \ (n \to \infty)] \\ &= 0 + 0 = 0,\end{aligned}$$

was $X_n + Y_n \to X + Y$ $f.s.$ impliziert.

14. Siehe z.B. Satz 11.4 in Bauer (1992).
15. Mit dem Satz von Fubini (siehe z.B. Korollar 23.7 in Bauer (1992)) sieht man diese Beziehung wie folgt ein:

$$\begin{aligned}\int_0^\infty \mathbf{P}[Z > t]\, dt &= \int_0^\infty \int_\Omega 1_{[Z>t]}(\omega)\, \mathbf{P}(d\omega)\, dt \\ &\stackrel{Fubini}{=} \int_\Omega \int_0^\infty 1_{[Z>t]}(\omega)\, dt\, \mathbf{P}(d\omega) \\ &= \int_\Omega Z(\omega)\, \mathbf{P}(d\omega) = \mathbf{E}Z.\end{aligned}$$

16. Einen Beweis dieses Resultats und verwandter Aussagen findet man z.B. in Kapitel IV in Bauer (1991).

Anmerkungen zu Kapitel 6

1. Die Formeln von Panjer werden z.B. in Mack (2002) beschrieben.
2. Einen Beweis findet man z.B. in Witting (1985), dort als Beweis zu Korollar 1.44.
3. Dies folgt z.B. aus Satz 1.19.1 in Gänssler und Stute (1977).
4. Für reelle Zufallsvariablen Z_n bzw. Z mit Verteilungsfunktion F_n bzw. F schreiben wir

$$Z_n \to^{\mathcal{D}} Z,$$

falls in jedem Stetigkeitspunkt x von F gilt:

$$F_n(x) \to F(x) \quad (n \to \infty).$$

Mit dieser Notation besagt Satz 5.12

$$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right) \to^{\mathcal{D}} N(0,1) - \text{verteilte Zufallsvariable.}$$

Wegen

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \to \sigma^2 \quad f.s.$$

(vgl. Abschnitt 6.2) kann man mit Hilfe des Satzes von Slutsky folgern

$$\frac{\sqrt{n}}{S}\left(\frac{1}{n}\sum_{i=1}^{n}X_i - \mu\right) \to^{\mathscr{D}} N(0,1) - \text{verteilte Zufallsvariable.}$$

5. Siehe z.B. Abschnitt 2.1.3 in Witting (1985).
6. Siehe z.B. Satz 3.2.2 in Gänssler und Stute (1977).
7. Einen Beweis von Satz 6.3 findet man z.B. in Abschnitt 10.2 in Gänssler und Stute (1977).
8. Einen Beweis von Satz 6.4 findet man z.B. in Witting und Müller-Funk (1995), siehe dort Beweis von Satz 5.130.
9. Siehe z.B. Satz 6.55 in Witting und Müller-Funk (1995).
10. Siehe z.B. Abschnitt 3.8 in Lehn und Wegmann (2000).

Literaturverzeichnis

1. Amaratunga, D. und Cabreva, J. (2003). *Microarray and Protein Array Data.* Wiley Series in Probability and Statistics, Wiley.
2. Bauer, H. (1991). *Wahrscheinlichkeitstheorie.* de Gruyter.
3. Bauer, H. (1992). *Maß- und Integrationstheorie.* de Gruyter.
4. Der, G., Batty, G. und Dear, I. (2006). Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *British Medical Journal* **333**, pp. 945-953.
5. Eisenhauer, N., Partsch, S., Parkinson, D. und Scheu, S. (2007). Invasion of a deciduous forest by earthworms: Changes in soil chemistry, microflora, microarthropods and vegetation. *Soil Biology & Biochemistry* **39**, pp. 1099–1110.
6. Ebrahim, S., May, M., Ben Shlomo, Y., McCarron, P., Frankel, S., Yarnell, J. und Davey Smith, G. (2002). Sexual intercourse and risk of ischaemic stroke and coronary heart disease: the Caerphilly study. *J. Epidemial Community Health* **56**, pp. 99–102.
7. Franke, J., Härdle, W. and Hafner, C. (2007). *Statistics of Financial Markets*. Springer.
8. Gadhia, P., et al. (2003). A preliminary study to assess possible chromosomal damage among users of digital mobile phones. *Electromagnetic Biology and Medicine* **22**, pp. 149–159.
9. Gänssler, P. und Stute, W. (1977). *Wahrscheinlichkeitstheorie.* Springer.
10. Grafarend, E. (2003). *Geodesy - the challenge of the 3rd millenium.* Springer.
11. Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. (2002). *A distribution–free theory of nonparametric regression.* Springer
12. Heart Protection Study Collaborative Group (2002). MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* **360**, pp. 23–33.
13. Heidrich, J., Wellmann, J., Hense, H.-W., Sieberte, E., Liese, A. D., Löwel, H. und Keil, U. (2003). Klassische Risikofaktoren für Herzinfarkt und Gesamtsterblichkeit in der Bevölkerung: 13-Jahres-Follow-up der MONICA Augsburg-Kohortenstudie. *Zeitschrift für Kardiologie* **92**, pp. 445–454.
14. Heuser, H. (1988). *Lehrbuch der Analysis, Teil 1.* B. G. Teubner.
15. Hildebrandt, H. (1998). *Pschyrembel Klinisches Wörterbuch.* Bearbeitet von der Wörterbuch-Redaktion des Verlages de Gruyter unter Leitung von H. Hildebrandt. de Gruyter.
16. Hull, J. (2006). *Optionen, Futures und andere Derivate.* Pearson.
17. Lehn, J. und Wegmann, H. (2000). *Einführung in die Statistik.* Teubner.
18. Mack, T. (2002). *Schadenversicherungsmathematik.* Schriftenreihe Angewandte Versicherungsmathematik.
19. Parzen, E. (1962). On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics* **33**, pp. 1065–1076.
20. Rogers, W. P. et al. (1986). Report of the Presidential Commission on the Space Shuttle Challenger Accident. Online erhältlich unter http://history.nasa.gov/rogersrep/genindex.htm

J. Eckle-Kohler, M. Kohler, *Eine Einführung in die Statistik und ihre Anwendungen,*
DOI 10.1007/978-3-642-15226-9, © Springer-Verlag Berlin Heidelberg 2011

21. Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics* **27**, pp. 832–837.
22. Roth, D. (2008). *Empirische Wahlforschung*. VS Verlag.
23. Schneider,R. (2002). Wettlauf um Virenkiller. *Die Zeit* **14**, pp. 38–39.
24. Singh et al. (2002) Effect of an Indo-Mediterranean diet on progression of coronary artery disease in high risk patients (Indo-Mediterranean Diet Heart Study): a randomised single-blind trial. *Lancet* **360**, pp. 1455–1461.
25. Tufte, E. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative.* Graphics Press, Chesire.
26. Weinstein et al. (2006). Surgical vs nonoperative treatment for lumbar disk herniation. The Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* **296**, pp. 2441-2450.
27. Witting, H. (1985). *Mathematische Statistik I.* Teubner.
28. Witting, H. und Müller-Funk, U. (1995). *Mathematische Statistik II.* Teubner.

Sachverzeichnis

J. Eckle-Kohler, M. Kohler, *Eine Einführung in die Statistik und ihre Anwendungen,*
DOI 10.1007/978-3-642-15226-9, © Springer-Verlag Berlin Heidelberg 2011